Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Procedia Computer Science 21 (2013) 472 – 478

Procedia Computer Science

The fifth International Symposium on Applications of Ad hoc and Sensor Networks (AASNET'13)

Selective Power Aware Data Management Approach for Wireless Sensor Networks

Hazem Morsy[a]*, Hesham El-Rewini[b]

[a] *Department of Computer Science, Southern Methodist University, Dallas, TX 75275, USA*
[b] *College of Engineering and Mines, University of North Dakota, Grand Forks, ND 58201, USA*

Abstract

The maturity of wireless technologies and the growing popularity of sensor technology in our daily life have created fascinating possibilities in the field of Wireless Sensor Networks (WSN). One of the most critical fields of WSN is how effective is the data collection model in such a network. The goal is to ensure data collection consistency and reliability while attempting to conserve the power consumption within the network. We are proposing a balanced data collection model that can be setup on random or controlled WSN deployment that can meet different levels of reliability requirements while balancing power consumption across the network. The proposed model utilizes an optimization heuristic to achieve a set of system-defined goals.

Keywords: Wireless Sensor Networks; mobile computing; wireless networks; Distributed systems; Data Structures; Optimization theory.

1. Introduction & Background

Wireless sensor networks (WSN) have emerged in recent years as an important technology with applications in many fields such as security, weather monitoring, agriculture, space explorations, emergency response and many more. In addition, the increased popularity and usage of wireless mobile devices (smart phones, tablets) accompanied by advancements in wireless technologies, data exchange and mobility support have enabled the integration and expansion of WSN in such fields. Such integration enabled coordinated infrastructure that allows location independent sensor data sharing in a dynamic manner. “The emergence of wireless sensor networks (WSNs) is essentially the latest trend of Moore's Law toward the miniaturization and ubiquity of computing devices” [1]. This evolution introduced new challenges stemming from the dependency on unreliable wireless connectivity, deployment challenges and dependency on limited power sources. Such challenges need to be

* Corresponding author: Hazem Morsy Tel.: +1-972-510-7623; fax: +1-972-348-9736.
E-mail address: hmorsy@smu.edu.

doi:10.1016/j.procs.2013.09.064

addressed in order for such merge of technologies to reach its intended potentials. Data collection is at the center of such challenges where the collected data needs to be properly aggregated, fused and routed to the final destination [2]. The activity of data sensing and data routing (especially when sensors has to execute routing logic) can take their toll on the sensor battery. This can lead to defragmented WSN and to some extent a disjointed senor network [3]. Basically, the life of a WSN is directly proportional to the life of its sensors. Losing a single sensor could result in the loss of the whole network if that sensor ends up being a critical node that is at the center of any data collection routes to the destination node (data sink). Hence, It is important for any dynamic WSN to collect its data in a reliable manner while conserving or balancing the power consumption within its sensors. In this paper we are proposing an adaptive data collection approach that attempts to balance power consumption while optimizing route selection to achieve system desired Service Level Agreements (SLAs).

Wireless sensor networks (WSN) were the result of recent advancements in wireless communications and electronics [4]. Arampatzis in [5] indicates that “the collaboration and synergy of sensing, processing, communication and actuation is the next step to exploit the inheritance of this new technology”. This has resulted in the ability to develop and deploy low cost wireless sensors across different types of environments [6]. In general, a WSN is made of a large number of battery-driven sensors that could be deployed in a random or controlled manner in the area of interest [7]. WSNs support two standard technologies for their interaction and data exchange activities: ZigBee and Bluetooth. “Both operate within the Industrial Scientific and Medical (ISM) band of 2.4 GHz, which provides license-free operations, huge spectrum allocation and worldwide compatibility” [12]. In general, a node in a WSN senses data from its surroundings and transmits it to more capable nodes, called data sinks. These data sinks aggregate and analyse sensed data to facilitate a designated function of the WSN [6]. Examples of these functions were discussed in [9] where WSNs can be utilized in security and military application, in auto response rocket launchers and in emergency disaster response. Furthermore, the authors in [11] add that WSN networks are oriented towards pervasive computing through the sensing and integration of the collected environmental data.

The possibilities and challenges offered in WSN explorations are widely recognized in both academia and industry [5]. One of the most critical design decisions in a WSN is how the sensed data is collected and routed. The data collection approach can have a very detrimental impact on the life span and effectiveness of the WSN. Wand and Liu in [2] categorized the stages of data collection in a WSN to three types: the deployment stage, the control message dissemination stage and the data delivery stage. Each stage has its characteristics and its challenges. Chen et al in [14] studied capacity challenges in WSN data collection stemming from the different types of collection scenarios. They proposed an asymptotic upper bound approach that can control this problem. In other research work like [3], data collection in WSN was studied with multiple sinks for large data collection. They defined an approximation model that attempts to minimize data collection latency with a constant-factor performance guarantee. Another approach proposed in [10] using network modelling to predict power consumption within the network and attempt to optimize it. The model is best suited for the spatial correlations and broadcast nature of communication within the wireless network. Finally, further models and approaches were studied in different research work like [15,16] to attempt to address the data collection challenges. For the most part, this work had focused on either maximizing amount of data collected, minimizing data latency or minimizing power consumption among the sensors. However, none of the work attempted to combine these goals and to balance the outcome of such conflicting goals.

Furthermore, any working data collection implementation has to rely on a stable and reliable Medium Access Control (MAC) layer. The MAC layer is responsible for accommodating data transmission by the different wireless sensor nodes that are sharing the limited wireless bandwidth. This layer manages nodes’ packet scheduling and transmission in a fair and effective manner while utilizing the physical link layer below it. One of the key fundamental properties of a WSN MAC layer is to minimize packet collision while maintaining an adaptable, scalable and energy efficient communication protocol. Demirkol et al stated that “other important attributes such as latency, throughput and bandwidth utilization may be secondary in sensor networks” [17]. This research work continued to survey different protocols including Sensor-MAC (S-MAC), WiseMAC, Traffic-Adaptive MAC and tree-based data MAC (DMAC). They concluded that each of the proposed techniques had their advantages and disadvantages and neither one provided a pure answer to addressing the key set of objectives of scalable, energy efficient and dynamic MAC layer. They add that cross layer integration between the MAC layer, link layer and routing layer can provide a more promising solution to these challenges and objectives. Furthermore, the authors in [18] focused on Quality of Service (QoS) based MAC that can be controlled to drive specific set of quality control parameters to address sensors communication with sink nodes. The quality control is parameters setup will be driven by the nature of the data (or packet) to be transmitted.

It is important to note that the work we are proposing in this paper assumes that there is a reliable, energy efficient MAC layer that will be utilized to execute the routing pattern defined by the proposed data collection approach. We will not be focusing on the nature of the MAC layer used here as we are focusing on the data collection protocol. The paper is organized in the following manner: Section 2 will cover the proposed WSN architecture and system components along with WSN states. Section 3 provides the details of the proposed data manager approach. Section 4 discusses the simulation model and the achieved results. Finally, we provide our concluding remarks in section 5.

2. Platform Components

The proposed architecture utilizes a set of sensors that are deployed in an uncontrolled crisis environment (such as earthquakes, floods or disaster sites). In such a deployment, it is desired not to have predefined sink nodes within the network. In such an environment, there are no guarantees that these sink nodes will be positioned in the proper way to handle the data collection from the rest of the WSN. Each sensor needs to only discover nearest neighbors that are within its communication range (this will be a key advantage when it comes to data collection scalability). The proposed data sink will be a mobile handheld device that subscribes to the WSN to receive the sensed data. This device is characterized as a portable device (like a smart phone, tablet or laptop) that is carried around the sensed field to collect data. This device will be referred to as the Sensor Data Manager (SDM). A SDM will have the needed software and APIs to enable it to act as a network moderator, data sink manager and sensor assignment coordinator. The idea in this architecture is that this on demand environment can be setup with minimal infrastructure requirements. The presence of such dynamic/mobile data sinks (SDMs) will eliminate the need to build or implement special sensors inside the uncontrolled field of deployment. The most critical function in an SDM is acting as the data collector to initiate or receive sensed data from the field. This function can be set up as an on demand request initiated by the SDM or it can an automated event where sensors send their data updates to all subscribing SDMs at a predefined frequency. Figure-1(a) shows a deployment example with an SDM subscribing to the WSN and requesting data that can be provided by sensor S from the network. We will use this example to describe the data collection approach. Before we get into the algorithm, we will need to describe the key phases of the network. The proposed WSN goes through three states that govern its operations. These states assume that the sensor nodes are equipped with the software and hardware needed to support the defined functions.

1- Deployment State: in this state, sensors are deployed in the location or environment that it needs to survey. This deployment can be short-term deployment where the sensors are used for a specific task for a limited time (such as emergency response in disaster situations) or it can be long-term deployment (like in precision agriculture). The deployment in both cases is assumed that it can only be random where there is no control on how the sensors are placed (other than just attempting to deploy them in close proximity within the desired area). In such deployment, we assume that this environment is unstable and predictable and changes within the environment can cause the sensors to shift location over time. Hence, it is important that the proposed data collection model can adapt to such unpredictable behavior. When the sensors are deployed, they begin to identify themselves within the WSN by sending a quick ping signal. All sensors within range of each other begin to establish their **Degree of Connectivity (DC)**. This is basically the number of nodes that a sensor can reach in a single hop. For example, In Figure-1(a), node 3 has a degree of connectivity of 4. Node 3 has direct access to nodes 2, 5 and 6 in addition to its own sensed data. Hence, node 3 can aggregate data from 4 different nodes and can provide 3 different routing options through its direct neighbors.

2- Sensing and Data Collection State: In this state, the sensed data is collected and routed to the nearest node where a SDM is directly connected. The SDM can initiate the data collection request to the nearest set of sensors and that request gets propagated through the network (this will be part of our future work). In this paper, we will be addressing the automated data collection model where sensed data are sent by each sensor at a desired frequency to the subscribing SDM. It is important to note that SDM will subscribe to more than one sensor within the network to minimize the impact of having a single point of failure where the sensor can rapidly be depleted of its power. Furthermore, since the SDM is a mobile device, it is expected that the SDM can change its location around the WSN in an active manner. It is important that the WSN data collection protocol can provide a dynamic subscription

process where the desired data can be collected from any where in the network to the SDM. Finally, Each sensor in this model will be equipped with the software component that can execute the defined data collection protocol.
3- Maintenance and Reconfiguration State: This is a refresh state that any of the WSN sensors can use adjust its statistics operating parameters. Parameters such as the sensor's degree of Connectivity and data collection frequency can be adjusted during this state. This will enable the WSN to adapt to the dynamic changes within the environment. It also serves as an optimization stage for the SDM to recalculate the their access points to the WSN (based on sensors reliability statistics). The adjustment process can be an automated process or a system administrated one. This state is triggered by a predefined event like a designated maintenance window that is done at a predetermined frequency. This may not be possible in a disaster environment where the only option is to set an automated update frequency process before deploying the sensors.

3. Selective Power Aware Data Collection

The selective power aware data collection approach is based on the minimum cost optimization heuristic with system parameters that enables the system to balance power consumption across the network while selecting best possible routes that has the highest chance of success in getting the data to the subscribing SDM(s) under the desired latency requirements. This can also be viewed as the data collection protocol ability to maximize the amount of data collected system according to the defined Service Level Agreement (SLA). The idea is to accommodate the different dimensions of the problem (SLA requirements, data delivery assurance, power consumption control) by mapping the problem into a network flow representation of the WSN. Figure-1(b) shows an example of the WSN and how it is mapped into a network flow. Each sensor is represented as a node in the graph with sensor S being the originating data source and SDM being the destination node. Each node will be able to build its local graph view of its nearest neighbors during the initial ping process at the deployment phase. This image will not refresh until the next maintenance or update phase. A key design factor to be considered is the frequency of these updates as they are considered an overhead activity. However, since this is a controlled factor, it can be set at a desired level that would be most suited for the deployed network. Part of the future work we are considering is to incorporate the latest node statistics to be part of the data collection packet. For the scope of this paper, we will be assuming a fixed maintenance or update window that is predefined during system setup. As each node builds its nearest neighbor segment of the graph it will be able to handle the decision making on the route selection for its part of the graph. Hence, the optimization approach here will be localized at each sub graph level rather than the global level. The edges in the graph represent connectivity between the nodes. The capacity on each edge will reflect the degree of connectivity on the receiving node. The cost on each edge will reflect the inverse of the battery level on the receiving node. For example, the edge from node 1 to 4 will have a capacity of 4 since node 4 has a degree of connectivity equal to 4 nodes (including itself). The cost on the edge will reflect the inverse of the battery power percentage on node 4 (node 4 has a battery percentage of 40% remaining). Hence, when applying the minimum cost function that combines the degree of connectivity (reliability of the route) and the cost function (selecting routes with higher power levels), the route selection will drive towards balancing power consumption while improving the reliability of data collection route. Figure-1(c) reflects a pure power optimization objective when routing and aggregating data from node S to the SDM. The selected route will aggregate data from the desired node S through nodes 1, 4, 8 until it reaches the designated SDM. These nodes exhibit the highest power availability levels at this point in time. As time progresses, this route will become less desirable as its battery levels will drop below other available routes. Over a period of time, the route selection process based on the minimum power consumption will end up balancing power across the different routes resulting in maximizing the life expectancy of the WSN. In this example, the primary objective was to balance power consumption. However, we can incorporate other objectives in the data collection optimization goal. For example, we can incorporate the SLA to maximize the amount of data selected while still maintaining an acceptable balanced power consumption levels. In this case, the data collection model will select routes with the most degree of connectivity if the delta between the edges' cost is within an acceptable threshold. For example, if the threshold is set to 35%, and if two routes were available for the next node visit, the algorithm will force the selection of the route with the highest degree of connectivity if the difference in power level between the two routes is less than 35%. Figure-1(d) shows an example where the algorithm uses the selected route in red (S1 through nodes 2, 3, 5, 8 to get to SDM). Hence, this approach is not a pure maximization approach but rather an intelligent route selection approach that attempts to maximize the amount of data collected as long as we are within the acceptable power balancing goals. This approach will still balance power consumption across the network but at a slower rate than the previous example. In this approach we will be using the maximum data collection route option until we are over the designated threshold, the algorithm will then shift to the next less desirable route (with less data collection opportunity but with better power consumption rate) until the power threshold difference between the maximum route and the lesser one shifts again towards the maximum route (or a third route is identified). The previous example had an acceptable threshold level of 0, and hence, the selected route will always be the least expensive route (in terms of power consumption). Finally, we can incorporate other objectives within the same algorithm. For example, we can incorporate data collection latency as the driving objective in addition to the power balancing approach. In this case, the capacity on each edge will reflect the distance to the SDM. The closer you are to the SDM, the higher your capacity will be. This will indicate proximity to the destination. There are different ways that can be used to determine the proximity of a sensor node to the SDM. Different techniques described in [13,15] can be incorporated to calculate the Degree of Proximity (DP), which will be reflected on each edge (instead of the degree of connectivity). In this case, the optimization problem of the formed network flow attempts to ensure traversal of WSN at the least amount of time while minimizing the cost of power consumption to achieve that goal. Hence, applying the same selective minimum cost approach to optimize against data latency can be applied as long as we are within the acceptable power consumption threshold between the different routes. The algorithm will still behave in a similar recursive manner attempting to balance these multi-dimension objectives while maintaining acceptable power consumption balance across the WSN. The simulation study conducted in this study utilized the latency objective as the second objective along with the power balance objective. Another key advantage of this modeling approach is that the dynamic nature of the WSN and multi-dimensional objectives can lead to different route possibilities when collecting sensed data. This makes a network flow representation of the WSN a more natural representation as it can dynamically reflect nodes availability (inadvertent movements or loss of power) by disconnecting and reconnecting the different vertices.

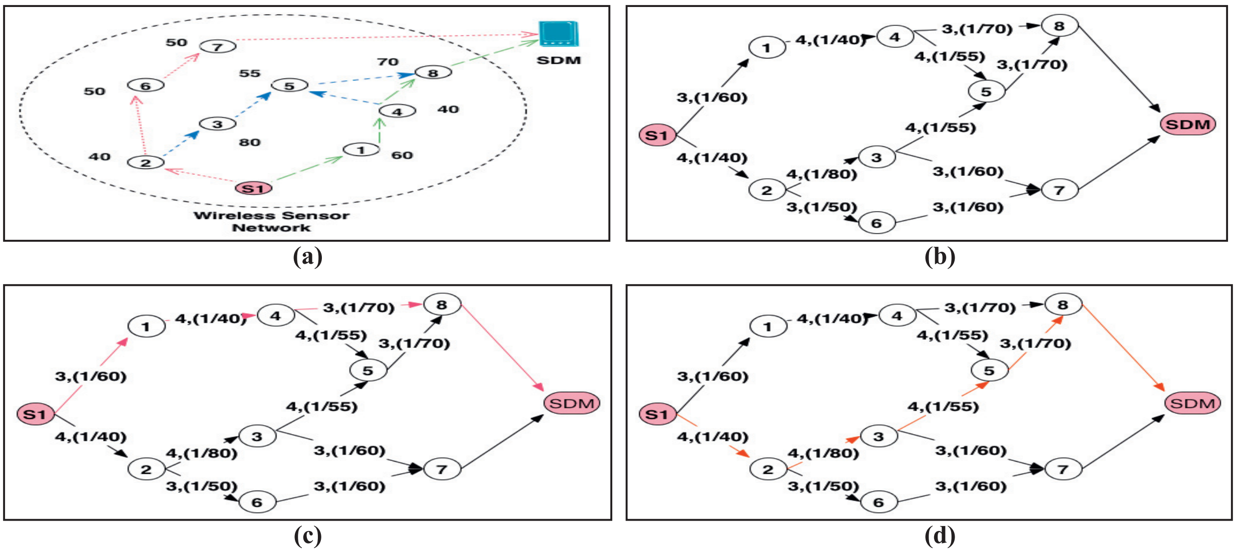
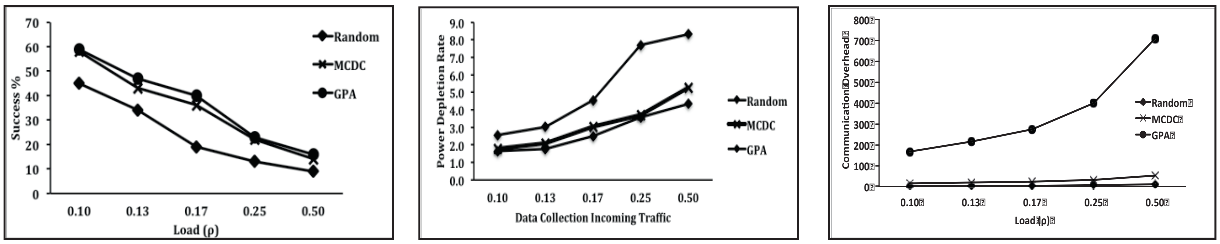


Fig.1 (a) Sample WSN layout with a subscribed SDM. (b) Demonstrates how the WSN example is converted to a network flow graph with capacity reflecting degree of connectivity and cost reflecting power consumption. (c) demonstrates route selection using the minimum cost approach with the power balance being the primary goal when collecting data from S1. (d) demonstrates route selection from S1 to SDM with the SLA of maximizing the amount of data collected while maintained an acceptable power consumption balance across the WSN.

4. Simulation Study

The simulation model was developed with SimJava v2.0 to implement evaluation experiments. This simulation platform will be used to evaluate the Minimum Cost Data Collection (MCDC) approach against two baseline approaches. This will help in verifying whether this approach can deliver on its projected promises. We compared the proposed data manager against two data collection standards. The first approach is a simple random data collection protocol to traverse and aggregate the data from a WSN. This is a simple baseline approach that does not require any data exchange or information sharing (minimal overhead). This will help us gauge the proposed MCDC in terms of its incurred additional overhead exchange against the baseline minimal overhead demonstrated by the Random protocol. The second data collection approach that we will compare with is a greedy approach that attempts to use the shortest path algorithm to collect the sensed data. In this approach, the protocol will select the shortest path to the SDM that ensures traversing the path of nodes with the highest battery level. This approach will be labeled as Greedy Power Aware (GPA) data collector. This greedy approach will provide the upper bound for the best possible power consumption across the WSN as it always tries to pick the best route with the most power level. This will serve as a benchmark to assess the quality of the MCDC power balance approximation algorithm when compared to the best greedy model. The simulation model evaluated a set of parameters that we will be presenting a sample of the simulation results. Data collection latency is used to measure the number of data requests that met the desired latency requirements over the total number of requests that were issued during a single simulation run. Figure-2(a) shows the outcome of this study against the data collection frequency requests. The GPA approach, which is a pure greedy data collection module, performed the best and was able to always find the shortest path (minimal latency). Our proposed MCDC performed relatively close to the optimal approach (especially during low traffic and extreme traffic). The next measure we evaluated in the simulation study is the average power consumption per sensor node in the network. Figure-2(b) presents the outcome of this measure in the simulation run. The GPA model will attempt to maximize the life span of the sensor battery by always selecting shortest paths with the highest power levels. MCDC performed well enough and behaved in a similar manner to the GPA greedy approach. Once again, this is achieved with a controlled heuristic execution that does not impose a heavy overhead on the WSN. It is important to note that GPA requires a global view of the WSN at any point in time. This means that scalability and execution overhead will be a major problem and may deem such implementation as unrealistic. This was very evident in the third measure we studied. The third measure is the communication and execution overhead needed to perform the different data collection algorithms during each simulation run. This measure is the sum of total data exchange runs between sensor nodes to update the WSN statistics within each node and to execute the route selection algorithm within each node. Figure-2(c) reflects the communication overhead cost from the simulation run. This measure shows the real advantage of using MCDC heuristic over the greedy optimal approach. The communication overhead of MCDC was resilient to increases in data collection frequencies. GPA on the other hand, had its communication overhead increase exponentially as traffic increased. This is a key advantage of using MCDC as it can achieve close to optimal results in terms of latency and power consumption while controlling the incurred overhead cost.



(a) (b) (c)

Fig. 2. (a) Percentage of data collection requests that met their Latency SLA requirements (b) Average power consumption per sensor against Data Collection traffic volume (c) Communication overhead cost across all nodes.

5. Concluding Remarks

The results presented in this paper are part of an effort to study data collection challenges in a wireless sensor networks. The proposed adaptive data collection approach based on the minimum cost heuristic has demonstrated strong potential and ability to perform close to optimal with controlled, low overhead costs compared to traditional greedy approaches. We have studied the behavior of the data collector under various conditions and environment setup. We intend to compare additional protocols within similar WSN setup that we proposed in this study. Furthermore, we are working on improving the proposed model when dealing with route and data redundancy.

References

[1] M. Bokare and M.A. Ralegaonkar, "Wireless Sensor Network: A Promising Approach for Distributed Sensing Tasks", Excel Journal of Engineering Technology and Management Science, vol. 1, 2012, p. 1-9."

[2] F. Wang and J. Liu, "Networked wireless sensor data collection: Issues, challenges, and approaches", Communications Surveys & Tutorials, IEEE, vol. 13, 2011, p. 673-687.

[3] S. Chen, M. Coolbeth, H. Dinh, Y.-A. Kim and B. Wang, "Data collection with multiple sinks in wireless sensor networks", Wireless Algorithms, Systems, and Applications, 2009, p. 284-294.

[4] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A survey on sensor networks", Communications magazine, IEEE, vol. 40, 2002, p. 102-114.

[5] T. Arampatzis, J. Lygeros and S. Manesis, "A survey of applications of wireless sensors and wireless sensor networks", Intelligent Control, 2005. Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation, 2005, p. 719-724.

[6] S. Ehsan and B. Hamdaoui, "A survey on energy-efficient routing techniques with QoS assurances for wireless multimedia sensor networks", Communications Surveys & Tutorials, IEEE, vol. 14, 2012, p. 265-278.

[7] L. Doherty, J. Simon and T. Watteyne, "Wireless Sensor Network Challenges and Solutions", Microwave Journal, vol. 55, 2012, p. 22-34.

[8] C.F. García-Hernández, P.H. Ibarguengoytia-Gonzalez, J. García-Hernández and J.A. Pérez-Díaz, "Wireless sensor networks and applications: a survey", IJCSNS International Journal of Computer Science and Network Security, vol. 7, 2007, p. 264-273.

[9] P. Baronti, P. Pillai, V.W. Chook, S. Chessa, A. Gotta and Y.F. Hu, "Wireless sensor networks: A survey on the state of the art and the 802.15. 4 and ZigBee standards", Computer communications, vol. 30, 2007, p. 1655-1695.

[10] P.J. del Cid, N. Matthys, D. Hughes, S. Michiels and W. Joosen, "Resource management middleware to support self managing wireless sensor networks", Self-Adaptive and Self-Organizing Systems Workshop (SASOW), 2010 Fourth IEEE International Conference on, 2010, p. 251-255.

[11] L. Zhaohua and G. Mingjun, "Survey on network lifetime research for wireless sensor networks", Broadband Network & Multimedia Technology, 2009. IC-BNMT'09. 2nd IEEE International Conference on, 2009, p. 899-902.

[12] L. Mainetti, L. Patrono and A. Vilei, "Evolution of wireless sensor networks towards the internet of things: A survey", Software, Telecommunications and Computer Networks (SoftCOM), 2011 19th International Conference on, 2011, p. 1-6.

[13] S. Chen, Y. Wang, X.-Y. Li and X. Shi, "Order-optimal data collection in wireless sensor networks: delay and capacity", Sensor, Mesh and Ad Hoc Communications and Networks, 2009. SECON'09. 6th Annual IEEE Communications Society Conference on, 2009, p. 1-9.

[14] L. Wang and A. Deshpande, "Predictive modeling-based data collection in wireless sensor networks", Wireless Sensor Networks, 2008, p. 34-51.

[15] A. Awan, S. Jagannathan and A.Y. Grama, "Scalable Data Collection in Dense Wireless Sensor Networks", 2007.

[16] L. Zhang, Q. Ye, J. Cheng, H. Jiang, Y. Wang, R. Zhou and P. Zhao, "Fault-tolerant Scheduling for Data Collection in Wireless Sensor Networks".

[17] I. Demirkol, C. Ersoy and F. Alagoz, "MAC protocols for wireless sensor networks: a survey", *Communications Magazine, IEEE*, vol. 44, 2006, p. 115-121.

[18] M.A. Yigitel, O.D. Incel and C. Ersoy, "QoS-aware MAC protocols for wireless sensor networks: A survey", *Computer Networks*, vol. 55, Jun. 2011, p. 1982-2004.